R E P O R T R E S U M E S

ED 011 234 SE 000 034

ASSESSING THE PROGRESS OF EDUCATION IN SCIENCE.

BY- TYLER, RALPH W.

PUB DATE SEP 66

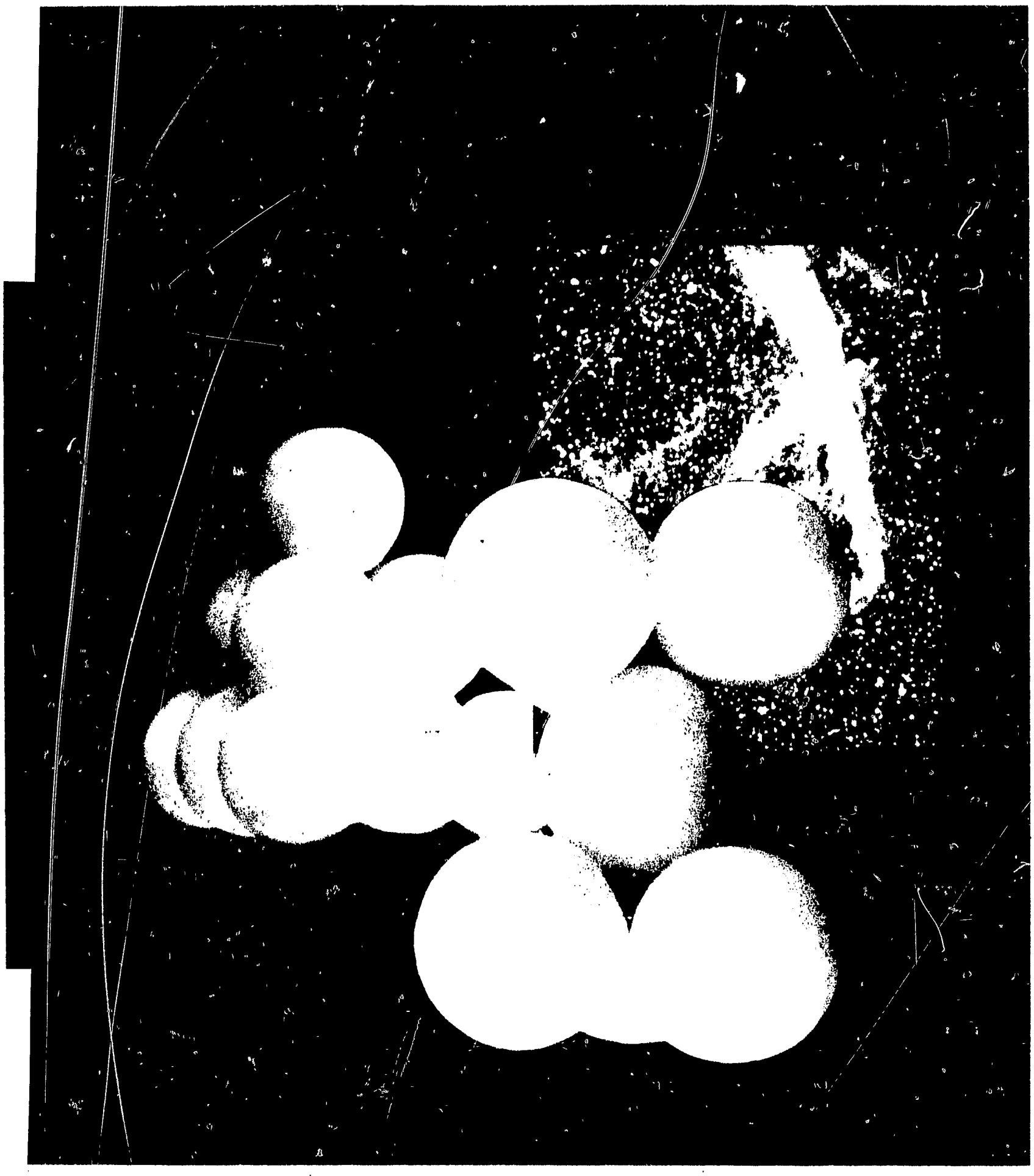
EDRS PRICE MF-$0.09 HC-$0.32 8P.

DESCRIPTORS- *EVALUATION, *EDUCATIONAL OBJECTIVES, *SCIENTIFIC LITERACY, *SCIENCE EDUCATION, ADULT EDUCATION, ATTITUDES, ELEMENTARY SCHOOL SCIENCE, SCIENTIFIC ATTITUDES, SECONDARY SCHOOL SCIENCE, COLLEGE SCIENCE,

THE PROGRAM IS DESIGNED TO ASSESS UNDERSTANDING OF SCIENCE AND THE SCIENTIFIC ENTERPRISE OF PERSONS IN THE UNITED STATES. INDIVIDUALS WILL BE RANDOMLY SELECTED FROM 192 POPULATIONS. SUBGROUPS WILL BE ESTABLISHED ON THE BASIS OF GEOGRAPHIC REGION, RURAL-URBAN-SUBURBAN DIFFERENCES, SOCIOECONOMIC CLASS, AGE, AND SEX. OBJECTIVES FOR SCIENCE EDUCATION HAVE BEEN DEFINED AND TEST EXERCISES BASED ON THE OBJECTIVES WERE CONSTRUCTED DURING THE SUMMER OF 1966. THE EXERCISES ON A TRIAL BASIS FOR FALL AND WINTER 1966-1967. INTERVIEWS AND OBSERVATIONAL PROCEDURES WILL ALSO BE USED TO OBTAIN INFORMATION ABOUT INTERESTS, HABITS, AND PRACTICES. PERCENTAGES OF RESPONSES OF TOTAL POPULATIONS WILL BE USED TO REPORT FINDINGS. FUTURE EVALUATIONS WILL BE MADE AND COMPARED WITH PREVIOUS RESULTS. THIS ARTICLE IS PUBLISHED IN "THE SCIENCE TEACHER," VOLUME 33, NUMBER 6, SEPTEMBER 1966. (AG)

VOLUME 33, NUMBER 6 · SEPTEMBER 1966

THE SCIENCE TEACHER



THE SCIENCE TEACHER

Journal of the National Science Teachers Association

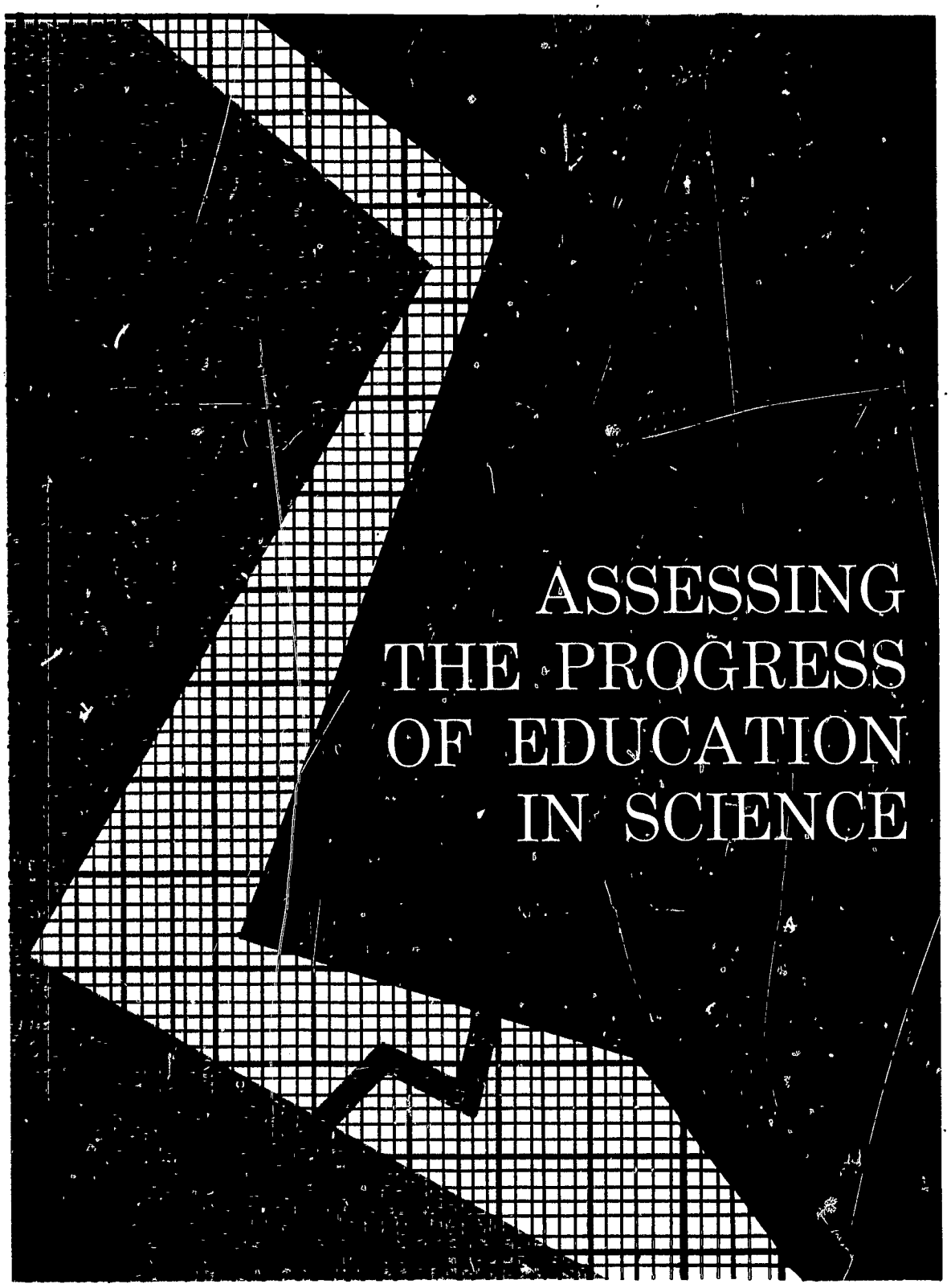
Volume 33, Number 6 • September 1966

COVER

Molecular models for phosphorus and an ethane-like structure as described by L. Carroll King in "Molecular Architecture" beginning on page 27.

Published monthly, September through May. Editorial and executive offices, 1201 Sixteenth Street, N.W., Washington, D.C. 20036. Second-class postage paid at Washington, D.C.

NSTA membership rates: Regular $8; Comprehensive $15; Student (university) $4; Life $200 (payable in ten annual installments, or $180 if paid in three years or less). Library and school subscriptions, $10; single copies of *TST*, $1.



RALPH W. TYLER

Director, Center for Advanced Study in the Behavioral Sciences
Stanford, California

WHILE general articles have described the national education assessment project, science teachers will be interested not only in the overall plans and progress of the program but more particularly with the efforts in the field of science education. This is a report of progress through the summer of 1966.[1]

The assessment project should be clearly distinguished from other useful educational appraisals that are commonly conducted in American schools. Probably the most frequent type of evaluation is to appraise the achievement of individual students. This is usually done with several purposes in mind. It may furnish a further incentive for students to study, because they know they will be tested. It may be used as one of the factors in grading or promoting students. It provides information that can be used by the student and counselor in planning for further education, and it often furnishes one of the bases for awarding scholarships.

A second use of evaluation is to diagnose the learning difficulties of an individual student or an entire class to provide information helpful in planning subsequent teaching. A third use of evaluation is to appraise the educational effectiveness of a curriculum, of a course, of instructional materials and procedures, and of administrative and organizational arrangements.

Each of these kinds of evaluation is

[1] For further background information, views of certain individuals, and statements by the Council of Chief State School Officers, the ASCD Executive Committee, and the AASA, see *National Educational Assessment: Pro and Con*, published by the National Education Association and the American Association of School Administrators. Stock No. 051-02094. 56 pp. $1.

an essential part of the processes of teaching or administration. Teachers and administrators are using evaluation of one sort or another as one of their normal procedures. The information gained from these appraisals is focused upon individual students' efforts, class performance, or the effectiveness of the plans, materials, and procedures used by the teacher, the school, or the school system.

There is another purpose for evaluation which is becoming increasingly important as our society becomes more complex and anonymous and as education becomes necessary for almost everyone in order to participate in the common life. In a democracy policies and actions depend heavily upon public understanding of the current conditions and the problems. Increasingly, the public in a community or state identifies its needs and approves proposed policies and actions against the background of understanding of the national and regional situation. Local and state policies and programs in the field of health, for example, have been sharply influenced by the national and regional information which is widely available regarding the age distribution of the population, the incidence of illness among different sectors of the population, and the availability of health services.

In the field of education, the public is often confused about the progress being made and the problems encountered because we do not have comprehensive data on the educational levels of various sectors of our population. We have reports on numbers of schools, buildings, teachers, and pupils, and about the moneys expended, but we do not have sound and adequate information on educational results. Because dependable data are not available, personal views, distorted reports, and journalistic impressions are the sources of public opinion, and the schools are frequently attacked and frequently defended without the necessary evidence to support either claim. The background of public opinion is an important factor in determining the moral and financial support given to educational policies and practices. Sound opinion requires valid data about the progress of American education.

This is the purpose of the assessment project. This type of evaluation is not focused upon individual students, classrooms, schools, or school systems, but furnishes overall information about the educational attainments of large numbers of people. Initially, the assessment will be made of the educational levels of four age groups—nine-year-olds, thirteen-year-olds, seventeen-year-olds, and adults.

RECOGNIZING the increasing need for public information of this sort, Carnegie Corporation of New York, a private foundation, in 1964 appointed an Exploratory Committee on Assessing the Progress of Education. I was asked to serve as chairman. The committee's assignment is to confer with teachers, administrators, school board members, and others concerned with education to get advice on the way in which such a project may be constructively helpful to the schools and avoid possible injuries. The committee is also charged with the development and tryout of instruments and procedures for assessing the progress of education. In 1966, the Fund for the Advancement of Education joined in the support of the project.

The discussions with administrators, curriculum specialists, teachers, and school board members clearly recommended that the initial assessment, though it could not cover everything, should include more than the 3 R's, and as rapidly as possible all the important educational tasks of the modern school should be covered. In harmony with this suggestion, for the initial tryout, instruments are being constructed by four leading test development agencies in the fields of reading and the language arts, science, mathematics, social studies, citizenship, fine arts and vocational education. In subsequent years, other important areas will be included. Educational Testing Service is the contractor responsible for the construction of the assessment instruments in the field of science.

THE contractor was responsible for bringing together science teachers, curriculum specialists, and scientists to formulate statements of the objectives which they believe faithfully reflect the contributions of science to the education of students and which the schools are seriously seeking to attain. It is important for an assessment that is taken seriously by the public to represent authentic science and to deal with learning which science teachers are earnestly seeking to develop. This should help to eliminate the criticism frequently encountered with current tests in which some item is attacked by a scientist as being "shoddy" science or criticized by school people as not in the curriculum.

The contractor was also required to clarify the meaning of each objective stated by including with it prototype exercises which, in the opinion of the science teachers and scientists, give students an opportunity to demonstrate the behavior implied by the objectives. This requirement was established to assure that the statements of objectives would not be so vague as to make their interpretation a matter of widely diverging opinions. Each objective is defined in concrete terms by the exercises which the science panel included as prototypes for the assessment.

The panel of science teachers, curriculum specialists, and scientists that were assembled by Educational Testing Service to formulate and define objectives of science education included: [2]

J. W. Buchta, executive secretary, American Association of Physics Teachers

Robert H. Carleton, executive secretary, National Science Teachers Association

Judson B. Cross, Physical Science Study Committee, Educational Services Incorporated

Charles E. Erickson, School of Chemistry, Rutgers University

[2] NSTA EXECUTIVE SECRETARY'S NOTE: To our knowledge, at least one additional panel comparable in diversity and stature of personnel was similarly assembled to serve in an advisory capacity in the area of science education during development test construction work by another contracting agency. Assurance has also been given that advisory and consultant assistance will be sought at other times as the assessment project moves ahead.—RHC

Frederick L. Ferris, Jr., Geology Department, Princeton University
R. Buckminster Fuller, Carbondale, Illinois
Glenn W. Giddings, General Electric Management Research and Development Institute
Bernardo F. Grossling, Geological Survey, U. S. Department of the Interior
John H. Marean, science consultant, State of Nevada
L. S. McClung, chairman, Department of Bacteriology, Indiana University
Floyd V. Moneghan, Department of Natural Science, Michigan State University
E. Duer Reeves, president, Reeves Associates, Incorporated, Summit, New Jersey

The professional staff from ETS were: Frank J. Fornoff, chairman, Science Test Development; Raymond E. Thompson, associate examiner; and William Kastrinos, associate examiner.

FOUR major objectives of science education were stated, delineated, and exemplified in the panel report. These four objectives are that students should come to:

I. Know the fundamental facts and principles of science.
II. Possess the abilities and skills needed to engage in the processes of science.
III. Understand the investigative nature of science.
IV. Have attitudes about and appreciations of scientists, science, and the consequences of science that stem from adequate understanding.

The delineation of I. consisted of a list of 26 topics in connection with which students of science can be expected to know fundamental facts and principles. Ten abilities and skills were outlined as necessary to engage in the processes of science, namely: A. Ability to identify and define a scientific problem; B. Ability to suggest or recognize a scientific hypothesis; C. Ability to propose or select validating procedures both logical and empirical; D. Ability to obtain requisite data; E. Ability to interpret data; F. Ability to check the logical consistency of hypotheses with relevant laws, facts, observations, or experiments; G. Ability to reason quantitatively and symbolically; H. Ability to distinguish among fact, hypothesis, and opinion, the relevant from the irrelevant, and the model from the observations the model was devised to describe; I. Ability to read scientific materials critically; and J. Ability to employ scientific laws and principles in familiar or unfamiliar situations.

Understanding the investigative nature of science was analyzed as:

A. Scientific knowledge develops from observations and experiments and the interpretation of the observations and the experimental results. Such observations and experiments are subject to critical examination and to repetition.
B. Observations are generalized in laws.
C. Laws are generalized in terms of theories.
D. Some questions are amenable to scientific inquiry, and others are not.
E. Measurement is an important feature of science because the formulation as well as the establishment of laws is facilitated through the development of quantitative distinctions. Measurements are inherently only approximate and are progressively inclusive and incisive.
F. Science is not, and will probably never be, a finished enterprise.

The attitudes and appreciations to be expected of one educated in science were also listed:

A. Recognize the distinction between science and its applications.
B. Have accurate attitudes about scientists.
C. Understand the relationship between science and misconceptions or superstitions.
D. Be ready and willing knowingly to apply and utilize basic scientific principles and approaches, where appropriate in everyday living.
E. Be independently curious about and participate in scientific activities.

The prototype exercises that were developed to define these objectives more concretely included different ones for the several age groups. For example, under *Knowing fundamental facts and principles,* nine-year-olds could be expected to know that green plants have the ability to produce food that animals use. Under *Possessing the abilities and skills,* a thirteen-year-old, when shown a film-clip of a walk in the woods, should be able to identify three phenomena that he could profitably investigate through a series of observations, measurements, and/or experiments. Under *Understanding the investigative nature of science,* seventeen-year-olds should be able to answer this:

"All of the following questions are amenable to scientific inquiry except:
(A) How does the motion of a body change under the action of a force?
(B) What metal alloy can be used to make the strongest nails?
(C) How can the most efficient heat engine be made?
(D) What proportion of the federal budget should be spent for space exploration?
(E) What colors of paint reflect the most light?"

Under *Attitudes and appreciations,* adults could be asked: "How often have you talked about current events in science with your friends? Which science magazines do you read? (For each one mentioned) How frequently do you read it?"

The above are only a few illustrations. The panel report outlines prototype exercises for each sub-objective and for each of the four age levels. This list of objectives, along with the lists from other fields and together with the prototype exercises which help to define them, have been reviewed by a series of panels of public-spirited citizens living in various parts of the country, in cities, towns, and villages. Each panel spent two days reviewing the material and making a judgment about each objective in terms of the questions: "Is this something important for people to learn today? Is it something I would like to have my children learn?"

This process resulted in very little revision of the original listing of objectives, but it should help to eliminate the possible criticism of the assessment as including unimportant matters or technical trivia.

AFTER the lists of objectives and prototype exercises were developed and reviewed, the contractors were asked to proceed with the construction of a large and representative sample of exercises for each objective and for each of the four age groups. This task was begun early in March 1966 and should be completed before the end of the summer. The exercises from the four contractors will then be assembled in convenient groupings for tryouts during the fall and winter of 1966-67.

A national assessment to identify kinds of progress being made in education and problems and difficulties arising will not be very meaningful unless separate measures are obtained for populations within the total country which vary among themselves and thus present different degrees and kinds of progress and different problems to be solved. The particular populations that need to be treated separately may change over the years ahead but for some time, age, sex, socioeconomic status, geographic location, and rural-urban-suburban differences will probably be significant. Hence, the present plan is to assess a probability sample for each of 192 populations defined by the following subdivisions: boys and girls, four geographic regions, four age groups (nine, thirteen, seventeen, and adult), three divisions by urban, suburban, rural classifications, and two socioeconomic levels.

The fact that populations are to be assessed and not individuals makes it possible to extend the sampling of exercises far beyond that of an individual test in which each person takes it all. It may be that a comprehensive assessment would require so many exercises that if it were to be taken by one person he would need 10 hours or more to complete them. With a population sample, 20 persons, each spending 30 minutes, would together take all the exercises. In this case, a population of 10,000 persons would furnish a sample of 500 for each of the assessment exercises, and no one would have given more than 30 minutes of his time. Assuming that an assessment would be made every 3 to 5 years in order to ascertain the kinds of progress taking place, it is very unlikely that many of those individuals who participated in the earlier assessments would be involved in any of the subsequent ones. Hence, from the point of view of the child or adult, no serious demand would be made on his time. Furthermore, it is unlikely that the children taking exercises in later years would be drawn from the same classrooms as the earlier ones. Therefore, the demands made upon a teacher in releasing a child for half-an-hour will be minimal. The assessment, though costly, should be feasible and involve little or no inconvenience to individuals or to schools.

Since the assessment does not require that all participants be in classes, the exercises to be used are not limited to the usual test items. Interviews and observational procedures are also to be employed to furnish information about interests, habits, and practices that have been learned. Because school objectives commonly include these areas, it is necessary to see that some assessment is made of the levels of attainment.

The assessment exercises will differ from current achievement tests in another important respect. An achievement test seeks to measure individual differences among pupils taking the test. Hence, the items of the test are concentrated on those which differentiate among the children. Exercises which all or nearly all can do, as well as those which only a very few can do, are eliminated because these do not give much discrimination. But, for the purposes of assessing the progress of education, we need to know what all or almost all of the children are learning and what the most advanced are learning, as well as what is being learned by the middle or "average" children. To devise exercises of this sort will be a new venture for test constructors. They are required to develop exercises at each age level in which approximately one-third represent achievements characteristic of most of those at that age level, one-third represent achievement characteristic of about half of those at that age level, and one-third represent the achievements characteristic of the most advanced, that is, the top 10 percent, of that age level.

TO summarize the educational attainments of these several populations it is not necessary to compute test scores. Instead, the following sorts of things would be reported:

For the sample of thirteen-year-old boys of higher socioecenomic status from large cities of the Northeast region it was found that:

91 percent knew two-thirds of the following important ingredients in a person's diet . . .

52 percent could plan an appropriate experiment for testing hypotheses like the following . . .

68 percent answered this question correctly: "Which of the following areas of scientific inquiry has been completely investigated and is thoroughly understood?

A. Electricity
B. Weather
C. Gravity
D. Heredity
E. None of the above"

57 percent indicated that they did not believe in any of the following superstitions. . . .

This mode of reporting avoids comparisons among individuals, schools, or school systems, and the meaning is more clearly understood by the layman than standard scores or other indices based on relative performance.

It is anticipated that the assessment would be in charge of a commission of highly respected citizens. They and the Commission staff would prepare reports of the findings of the assessment, much as we now obtain reports of the findings every 10 years of the decennial census. These reports would be available to all people interested in education, providing them in this way with significant and helpful information on what has been learned by each of the 192 populations. In subsequent years, the progress made by each of these populations since the preceding assessment would also be reported. # #